2 STORIES AND STYLES IN TWO MOLECULAR BIOLOGY REVIEW ARTICLES

GREG MYERS

The comments on the writing of review articles scattered in library journals and handbooks for writers of scientific articles stress the practical importance of reviews for the scientific community: pointing out that they collect, select, order, and interpret the huge outpouring of scientific reports, putting relevant findings and generalizations in a form useful to researchers outside the immediate group working on a prob1em.[1] But the comments also betray a certain uneasiness about the lack of originality in the genre, if only by insisting again and again on this originality.

> A scientific paper worth submitting to a journal must describe previously unpublished work. A review article will, of course, discuss previously published scientific work; its originality lies in the discriminating selection of material for comment and in the author's assessment of the current state of research on the topic under review. (O'Connor and Woodford, p. 4)

This issue is a problem because the one characteristic that the handbooks agree defines a review article—which can be interpretive or merely bibliographical, short or long, popular or specialized, in a review journal or a report journal—is that it does *not* report original work. So what does it do?

I will argue that the writer of a review shapes the literature of a field into a story in order to enlist the support of readers to continue that story. At any moment in the development of a field, the past has a canonical shape, recorded in the historical introductions of textbooks, in citations of "classic" articles, in eponymous terms. But the present is still a scattering of articles reporting various results with various methods aimed at various immediate problems. That's why classic research papers (see, for instance, those collected in Taylor) are often so hard to relate to the discoveries with which they are now associated; they are phrased in terms of immediate problems, while we understand the discovery in terms of a history leading to current work. The review selects from these papers, juxtaposes them, and puts them in a narrative that holds them together, a narrative with actors and events but still without an ending. It draws the reader into the writer's view of what has happened, and by ordering the recent past, suggests what can be done next.

The other crucial rhetorical feature of a review is its style, particularly the persona the writer presents in relation to the readers outside his or her specialty that he or she wants to influence. The author may want readers to accept the author's claim or use it in their own work, as in a research report. But a review may have a more subtle persuasive goal; it may define and present the whole topic so that readers see it in a certain way, as moving in a certain direction, so that it relates to them. And if they see themselves as part of a line ofwork, if they see the problem as their problem, they contribute to the power of its proponents (see Latour). But the influence does not all travel one way, from writer to audience.[2] The discovery of this broad audience is also a rediscovery of the topic. In a specialized research article, the writer can take for granted certain assumptions and methods, knowing that any competent specialist reader will also take these ideas and methods for granted, and will be looking for what is new, the claim. In a review article, it is just these assumptions and methods that must be brought out and put into an apparently logical order.[3] The writer can make sense of his or her field as a whole because he or she sees it from outside, with these readers, and has to ask the always risky question, "So what!"

Francis Crick suggests the importance ofthis two-way relation in his response to Horace Freeman Judson's interview question about a symposium paper Crick gave in 1957. The historical accounts of this period suggest Crick's paper had an influence on the development of research at the time, changing the way researchers saw current research (for instance, in his discussion of the Central Dogma) and enlisting support from experimentalists in their plans for future work (to test the Adaptor Hypothesis).[4] What is interesting about Crick's comments, whether one accepts them as historically accurate or not, is that he chooses to present himself as finding his ideas only in response to the need to state them to a symposium audience, to a general audience of biologists outside his usual circle.

> But you realize that what one was called upon to do for that symposium was to write a review article. To write a review article you had to put your ideas down on paper. You then express ideas which you hold but you didn't know you *held.* (Judson, 337)

The Rhetorical Situation

We can see the shaping of the field and the interaction with readers by close analysis of the stories and styles of reviews. I am going to compare two reviews on the same topic. One is James E. Darnell Jr.'s 1978 *Science* article, "Implications of RNA-RNA Splicing in Evolution of Eukaryotic Cells," which I take as fairly typical in its story and style. The other- Francis Crick's "Split Genes and RNA Splicing"—is atypical in many ways; it appeared in *Science* four months later. This article was written more than twenty years after the 1957 symposium paper, but Crick's reviews still have an important role in shaping the field.

My argument is that both articles show the rhetorical problems and purposes at work in review articles. Since Darnell's article is much more like most reviews, it might seem that I could illustrate my point with just one text. But my argument will be stronger if I can show that this explanation of the relation between form and function applies even to texts that would seem to be atypical. Jonathan Potter and Margaret Wetherell make this methodological point in their introduction to discourse analysis:

> If the proposed functional analysis is correct it ought to make sense of both the pattern regularly found in the data and the exceptions. That is, the analyst must predict that there will be special features of the organization of the exceptions which allow them to fulfill the required function in some other way. (69)

I am arguing that both the apparently odd text by Crick and the more typical text by Darnell can be explained as ways of telling a story about the past that shapes the future.

The form of Darnell's review presents the various results as offering a choice between two rival stories. But the style of Darnell's article plays down any sense that he is attempting to enlist support for his own view; the reader takes his view as the result of an apparently logical, impersonal process. Crick's article, on the other hand, seems to have no story at all. But gradually we see that he is focusing attention on the process of sorting out findings, rather than on the result; he is enlisting support, not for a claim, but for a way of formulating claims. For Crick's persuasive purpose, the impersonality of Darnell's article (and most other reviews) would be inappropriate. Instead he uses an informal style, as if inviting the reader to join a discussion among molecular geneticists.

One reason for these differences in style is that the two writers are famous for different things and thus have different rhetorical problems. Darnell has the rhetorical problem of an experimentalist using his lab's experimental results without wanting to seem to promote them. Crick has the rhetorical problem of any theoretician, sorting through the research on split genes when he himself hasn't done any of the research.

Darnell is one of the more heavily cited researchers working on RNA processing, the director of a large laboratory at Rockefeller University (which has a distinguished tradition in molecular biology), and author of a number of reviews and popularizations. An article that might serve as representative of the work of Darnell's lab is their contribution to the 1977 Cold Spring Harbor Symposium (the meeting where split genes were first discussed). In it they summarize several years of work that showed that in adenovirus 2, a DNA virus infecting human cells, the RNA that is first produced must be processed to produce the messenger RNA (mRNA) that codes for the protein. At the same symposium, other researchers announced the discovery of split genes. As we will see, Darnell knew of the discovery but planned his group's paper to stress a related but distinct issue. Whether Darnell's work is presented as part of the discovery of split genes or as another line of work depends on how reviews present it. Readers might see this *Science* article as a chance to hear an experimentalist relating a range of experimental reports to a highly speculative topic. But because he is known to be involved in the field, they might also be alert for any sign that he was using this format to promote his own experimental work.

Thanks in part to Watson's *The Double Helix* and to a number of histories, including a recent television movie, Crick is one of the most publicly known biologists of our time. His most famous papers— besides the 1953 article with James Watson that proposed the double-helical structure of DNA, and the 1961 article with a team at Cambridge that showed the triplet structure of the genetic code—include several reviews. These were usually written originally for symposia, conference keynote addresses, or lectures, and often, like the 1957 symposium paper I have mentioned, introduced large concepts which have since entered the textbooks. Crick did not, like Darnell, contribute experimental results to split genes research, so he could hardly be accused of plugging his own work. The rhetorical danger is that some readers might see him as playing around with other people's data, without offering anything of his own.

To understand Darnell's and Crick's contrasting approaches to a review on split genes, we need to try to reconstruct what the field would have looked like when they were writing in the late 1970s.[5] Both articles refer to recent dramatic developments in the field that are the reason for commissioning the reviews; Crick mentions the specific meeting where the discovery was announced.

> By the time of the annual Cold Spring Harbor Symposium, in the summer of 1977, it was clear that there was something very strange about the arrangement of the genes in several mammalian viruses, and for this reason it seemed highly likely that some

> chromosomal genes would also be in several pieces. This has since been found to be the case. (Crick 264)

At the conference to which Crick refers, groups of researchers from both MIT and the Cold Spring Harbor laboratory announced that parts of the genes on the adenovirus DNA were separated by strings of DNA that didn't appear on the messenger RNA (these intervening strings were later termed *introns* by Walter Gilbert). Previous work on cells without nuclei *(prokaryotes* such as bacteria) had led biologists to assume that the code on the DNA would correspond directly to the code on the mRNA and on the protein. The discovery with viruses (which use the cell's own mechanisms) suggested that cells with nuclei *(eukaryotes* such as those in yeasts, mice, rabbits, chickens, and humans) process the RNA in a much more complex way than had been thought.

Both Crick and Darnell were at this Cold Spring Harbor conference. Since we are interested in the persona Darnell chooses to project, it is important to ask how his own work on adenovirus RNA is related to the work he would review. Darnell gives this account of how he presented his work in relation to the discovery of splicing.

> I knew about the two splicing papers before they were presented at CSH that Friday (I think) night. We purposely did not go ahead to discuss any of that in our CSH paper but rather featured the point that all our evidence taken together compelled the view that large hnRNA [heterogenous nuclear RNA—large strands of various lengths found in the nucleus during infection] was the mRNA precursor. . . . This . . . is one of the reasons why the splicing evidence (in the form of EM [electron microscope] pictures) was so readily accepted as a biochemical fact. (Pers. com., 13 November 1987)

In this account, Darnell's work did not show split genes, but provided evidence that suggested a reason for splicing—to trim precursor molecules. (Some news articles, e.g., those of Sambrook and Schmeck, present Darnell's work in relation to the discovery in this way.) His decision not to refer to the other papers at the conference shows an awareness of the need to place his work effectively, so that it is neither subsumed in the new developments nor separated from them.

Starting in September 1977, reports began to appear, from researchers who had been at the Cold Spring Harbor conference, confirming that genes were split, not only in the viruses infecting eukaryotic cells, but in eukaryotic cells themselves. The data piled up as researchers, using the new sequencing techniques and the new recombinant DNA techniques, quickly found introns in a very wide array of eukaryotic genes. These reports came so fast that in 1977 and 1978 the situation was hardly stable from one month to the next; what was pure speculation in June was news in September and by February 1978 it was a piece of knowledge that could be indicated with just a note. Brief reviews of the adenovirus work appeared in the news columns of Science (Marx) and Nature (Sambrook) even before papers from the Cold Spring Harbor Symposium could appear in *Cell* and *PNAS.* In 1978 and 1979, the first large-scale review articles on what was now called "split genes" could already list several hundred articles—while apologizing for leaving so many out (see Abelson; Breathnach and Chambon).

One effect of all these reviews taken together was to create "split genes" as a topic. To see this effect, we have to distinguish the view of the field by outsiders such as science journalists (or myself), for whom the field is a series of events and distinct topics, from the view of those working in the field, for whom topics are defined by the daily work of research. No laboratory set out to study split genes; they set out to study RNA transcription units, or cell differentiation, or the generation of diversity in the immune system, or processing controls on genetic expression. And no lab continued to study "split genes" afterward; they study protein domains, or molecular evolution, or self-catalyzing RNA. One of the discoverers of splicing in adenovirus 2 RNA comments that the discovery, which is so important in popularizations and in reviews, soon became just a working fact for the researchers themselves. "The discovery of splicing was a singular event in a fairly large scientific field of research and it was so rapidly incorporated into the conceptual and experimental activities of the community that the fact of the discovery was soon ignored (Phillip Sharp, pers. com., 21 April 1987). The topic "split genes" arises as reviews and popularizations relate all these separate research programs around what can be called one phenomenon. This is not to say that the first researchers did not see the implications of their findings—they certainly did — but that it took the synthesizing activities of other researchers to make "split genes" a basic problem and to make the evolution of splicing a central issue.

These reviews have another social function, one that may seem unscientific to nonscientists: they encourage speculation. Crick and Darnell are searching for the constraints that further research, after the discovery, puts on possible scenarios for the evolution and function of the introns. The scenarios that remain after their sifting suggest further ways of designing and interpreting experiments that have led, for instance, to evidence for a once speculative notion, the "recruitment" of protein domains in primitive nucleic acids. As Crick puts it in his review, "This gap in our knowledge [concerning the origin of split genes] does not deter speculation, and for good reason, for such speculation may suggest interesting ideas and perhaps give us some general insight into the whole process" (268). Seen this way, speculation is not some imaginative flight tolerated in the last sentences of a research report; it is a crucial part of many reviews, for it focuses on the ending, the future work that will give the story its shape.

Plot and Story

The first problem for a writer like Darnell or Crick would be just reading and sorting out a great mass of papers, and then finding a way of parcelling out between forty and four hundred citations in some logical pattern. It can be a highly controversial matter who is mentioned, in what order, and what is said about them. It is not just that those who are omitted will be annoyed; those who are included will almost certainly find their work in a different context from that in which they themselves would put it. This is a problem of rhetoric, not just of mechanics. There is no template for the structure of a review article, no Introduction-Methods-Results-Discussion in which one can fill in the blanks. One handbook suggests this template is still useful.

> If you have previously written research papers and are now about to write your first review, it might help you conceptually if you visualize the review paper in terms of the research paper, as follows. Greatly expand the Introduction, delete the Materials and Methods (unless original data are being presented); delete the Results, and expand the Discussion. (Day 96)

But this is not very helpful advice, because the review writer cannot, as in a report, organize the introductory review to lead up to the work reported, or organize the concluding discussion to show its possible importance; one cannot make one's own work the focus.

Although review articles do not follow the Introduction-Methods-Results-Discussion format of research reports, there are certain regularities related to the social function of the genre. Since they must appeal to a broad audience, the introduction usually defines a topic and siresses the importance of recent work on it. The conclusion usually presents the prospects for research in the near future. In between the statements are arranged in some larger narrative. It may be useful, in analyzing this narrative, to introduce here a distinction drawn in literary criticism between plot and story. Plot is the order of events as presented in the text. So, for instance, in Poe's "The Purloined Letter," Dupin and his friend the narrator are talking in Dupin's rooms, the police chief comes in and describes a case, time passes, and later Dupin explains the case to his friend. *Story* is the supposed chronology of events behind the plot: the queen is having an affair, the minister picks up her letter, the police search his residence without success and appeal to Dupin for help, Dupin visits the residence twice, recovering the letter and substituting his own letter for it, and finally Dupin explains the case to his baffled friend. The plot of a review article is what gives it a surface organization, often a complex table of headings and subheadings; it follows the chronology of the reader's experience of reading. The story is the underlying narrative it aims to convey, with a different chronology following some actors – molecules, biologists, methods, views – constructed in the article. (Perhaps in some scientific discourse the underlying structure is not chronological – but as far as I know it always is in biology.) I make this distinction between plot and story to bring out a difference between Darnell's article and Crick's. Although the plots of the two articles are similar – they move from the discovery of split genes to the evolutionary implications – the underlying story of Darnell's work takes the organisms as its subject, while the story of Crick's article is about the scientists themselves.

The *plot* of Darnell's article is summarized in the introduction.

> Acceptance of the hypothesis that, in eukaryotes, "spliced" mRNA molecules are frequently formed from non-contiguous sequences raises several interrelated questions: Why, when, and how in evolution did the divided genes arise? What function is served today by having genes remain divided? (1257)

To simplify, the plot has to do with research and argument among scientists:

> Recent experimental results cause a shift in thinking.
> Researchers present two views of when divided genes arose.
> One can propose two ways divided genes aid evolution.
> Researchers present evolutionary reasons for the retention of divided genes.
> Research strategies will change as a result.

The *story* of Darnell's article involves a new account of evolution in which the common ancestor of prokaryotes and eukaryotes would have had a splicing system, and the prokaryotes would have evolved by eliminating it.

> This article explores the idea that the complex of biochemical reactions that result in mRNA formation is the chief evolutionary basis that sets eukaryotes apart from prokaryotes. Further, the key evolutionary step is the ability of eukaryotes to utilize contiguous information in DNA. (1257)

So the story has to do with eukaryotic organisms:

> Eukaryotes and prokaryotes arose from a common ancestor.
> Eukaryotes maintained split genes while prokaryotes eliminated them.
> Eukaryotes were able to use split genes to their advantage in evolution.
> Eukaryotes keep divided genes today.

The various parts of the plot, the various sections of the article, all contribute their bit to this underlying story, this new account. The implications of this story for the practices of biologists become clear only in the last paragraph. If eukaryotes didn't evolve from prokaryotes, and use different processing mechanisms, then experimenters can't assume prokaryotes are models for the eukaryotic cell; they have to do everything all over again with eukaryotic cells. To readers who had done all their work on prokaryotes, this could be seen as a threatening view. As we will see in discussing style, Darnell is careful to present it, not as his view, but as the logical outcome of the comparison of two interpretations of the available evidence.

A similar skimming of the headings of Crick's article suggests that the plot moves from the experimental data to broader and broader speculations, from sections on the problem in general, to the extent of genes, the kind of molecules affected, and the length of the introns, to the mechanism of splicing, to evolutionary and taxonomic questions. The basis of his plot is to start with statements of facts that seem quite scattered, presented as if in simple lists—lists, say, of which genes have been found to have introns, or of how many introns various genes have, or of which sequences have been determined to date. When one does see very broad general ideas emerging, they emerge in such an offhand way that one might overlook them. The lack of order indeed is the order; we seem to follow the wanderings of the author's train of thought, but we are being led from specifics to some large evaluations of existing hypotheses and comments on what to look out for in the future.

Crick too mentions the adenovirus findings, but he presents them as an event in the world of researchers. His story involves, not genes, but ways of thinking about genes. Instead of presenting two views that sum up the field, he admits in each section the confusion into which the field has been thrown and then attempts to reestablish order with some tentative generalizations, always with a tension between the need for some clarity and the need to leave possibilities open. This focus is apparent in both his opening with the event at Cold Spring Harbor (the audience astonished, the researchers eager to take on the new problem) and in his closing comments on researchers' reactions to the shift the field has undergone.

> There can be no denying that the discovery of splicing has given our ideas a good shake. It was of course already surmised that the primary RNA transcript would be processed in some way, but I do not share the view sometimes expressed that splicing is only a trivial extension of our previous ideas. I think that splicing will not only open up the whole topic of RNA processing, which had become somewhat bogged down before splicing was discovered, but in addition will lead us to new insights both in embryology and in evolution. What is remarkable is that the possibility of splicing had not at any time been seriously considered before it was forced upon us by the experimental facts. This was probably because, looking back, we can see that there was no experimental evidence to suggest that such a process might be taking place, at least for mRNA. Lacking evidence we had become overconfident in the generality of some of our basic ideas. *(270)*

The take home message about relying on prokaryotic models is much the same as that in Darnell's article. But Crick steps back to look at the change in perceptions required by the new results. He presents a story that could almost be a folk tale, beginning with false security, which tends to stasis ("the whole topic . . . had become somewhat bogged down")and then to the discovery forced upon us." Finally he reaffirms the significance of the event, while others doubt it. The problem for him is not just evaluating the evidence, but evaluating the response of various researchers to the evidence, figuring out how the community works. Though his earlier articles tend to end on an optimistic note, foreseeing rapid advances, the 1979 article I am discussing here ends by throwing some cold water, saying that the rapid progress on split genes should not obscure the difficulties with transcription, where "we badly need additional breakthroughs." We will see later how this story was interpreted by one journalist.

Analysis of plot and story focus our attention on the level of events. On the level of sentences, one way reviews relate the order of statements—the plot—to the order of narrated events—the story—is through a characteristic verb structure. Both Crick and Darnell follow the sort of verb sequence John Swales (in *Aspects of Article Introductions)* has found to be typical of article introductions. Darnell's article, for instance, begins with a present perfect verb, implying a series of events leading up to and including the present moment:

> For some years evidence has been accumulating that messenger RNA (mRNA) formation in eukaryotic cells is substantially different and more biochemically complex than in bacteria (1-4).

Then it continues with present-tense statements giving the knowledge that the new research takes as given:

> At the 5' terminus most eukaryotic mRNAs from yeast to man contain a modified methylated structure called a "cap."

The move to past tense at some point signals the story of the work of the researchers, a narrative of human events contingent on techniques, luck, genius, or institutional organization.

> This very unexpected conclusion came first from work on adenovirus mRNAs.

The return to present tense signals the new state of knowledge.

> The only primary RNA transcript that can be detected from the regions of the spliced late mRNA's embraces all the spliced regions as well as all the major mRNA regions.

There are two types of past tense in Darnell's article, the past of particular events in laboratories last year (as in the example here) and the past of evolutionary events "1.0 to 1.5 X 109 years ago." Both types of past tense signal stories, and the present tense signals a return to the level of present knowledge, the level of the organization of the paper. Crick also follows this basic sequence.

> I have spoken as if splicing only occurred in the processing of mRNA, but we already know that at least two other species of RNA are spliced. Indeed, one of the earliest discoveries was that some of the transfer RNA (tRNA) molecules in yeast are spliced, although their introns are fairly small (9, 10). More recently, two groups of investigators have isolated a crude enzyme preparation that will perform the operation in a test tube (11, 12). (265)

As with Darnell, the past tense signifies events in human history ("one of the earliest discoveries was"), the present tense signifies truths about nature ("two other species of RNA are spliced"), and the present perfect focuses attention on recent research that is directly relevant to the present ("recently . . . groups of investigators have isolated). Both sequences allow a two-level chronology, as if in a complex novel with frame tales, flashbacks, and flash forwards.

Besides this verb structure, there is a structuring feature of reviews so obvious that it may be overlooked: the disposition of footnotes. Citations are the point of a review article. Darnell's article has a reference for nearly every statement, except for some summary statements and hypotheses, 72 references in all, so that the pattern of the article is an alternation between the claims of the cited texts and the comments of the review, or the general statements of the rcvicw and the details in the cited texts. Although Crick cites more than a hundred articles (he even apologizes at the end that because of lack of space, "References have been kept to a bare minmum"[7]) his text is not organized around the references as Darnell's is. Most of the references Crick does give occur in densely packed paragraphs of survey between paragraphs that explore ideas with few references. So, for instance, in the important (and often cited) section on the possible ways split genes could evolve, he refers to only a few articles; most of the notes refer to personal communications, Cricks own articles, or his own comments on various points. For Darnell the citations support each step of a complex argument. For Crick they are just the firs: step from which he begins his own thinking.

Styles

The persuasive power of a review arises, not just from the apparent coherence of its story, but from its ability to enlist readers, to make them see their own work as part of this ongoing project. To do this, the article sets up a relation between the writer and the readcr, creating a persona for the writer and making some assumptions about the knowledge and responses of the reader. I am going to look at some striking stylistic differences between these two articles in terms of differences in rhetorical stance. Persona is complicated in both articles because the authors speak for several points of view. Darnell divides the field at the moment into two "views" that are compared in one impersonal voice; the reader watches the demonstration. Crick seems to present a more personal voice, but it would be more accurate to say that he presents several voices in contrasting styles; the reader is invited to participate in the discussion. Similarly, the syntax and cohesion in Darnell's article suggest the impersonality of the reader-writer relation, while everything in Cricks article suggests an interaction between people.

We can see the differences in persona in the opening sentences of these two articles. Darnell begins:

> For some years evidence has been accumulating that messenger RNA(mRNA) formation in eukaryotic cells is substantially different and more biochemically complex than in bacteria (1-4). (1257)

The author though he cites himself, remains in the background; the subject of the main clause, evidence, relates to what I have called his plot, while the subject of the subordinate clause, *mRNA formation* relates to what I have called his story. We are asked to read the article because impersonal evidence requires a change in impersonal theories.

Crick, as we might expect from what we have seen so far, has an unusual opening:

> In the last two years there has been a mini-revolution in molecular genetics. When I came to California, in September 1976, I had no idea that a typical gene (1) might be split into several pieces and I doubt if anybody else had. (264)

The first sentence makes the same moves as Darnell's opening to attract the attention of other biologists, mentioning a new development with broad implications. But the second is oddly personal, in its use of *I,* in its orientation in personal time (*When I came to California* rather than *For some years*), and in the offhand comment on the rest of the molecular genetics community. We are asked to read this article because an important theorist has had his ideas shaken up.

The persona Darnell presents is both assertive (because he has a claim to make) and carefully impersonal (because he does not want to present it as just his own claim). The phrasing throughout shows how important it is for Darnell, as for most scientific writers, to keep a carefully impersonal surface even where, or especially where, there are personal commitments and choices underneath. His account begins with the Ad2 work, summarized in some detail, and the Ad2 work begins with his own lab's work on transcription units (footnotes 9, 12, and 13).

> This very unexpected conclusion came first from work on adenovirus mRNAs. Late in adenovirus type *2* (AD2) infection a series of at least 13 individual mRNA molecules (8, 9) were found to contain sequences from non-contiguous sites on the adenovirus genome (10, 11). That each of these "mosaic" mRNA molecules comes about by RNA-RNA "splicing" or "ligation" was inferred from studies on the synthesis of Ad2 specific RNA in the nucleus of the infected cells (12, 13).

He does not say in the text that he is referring to work in his lab; he could just as well be writing about someone else, and one must comb the notes to see that he had any role in this research at all. There are, of course, no references to his personal experiences or responses, except for the one mention of a "startling" post-transcriptional event (which implies someone to be startled). Other people are omitted too; there are no names in the article (with one exception), so everyone is reduced to a footnote number.[8] The depersonalization is not just conventional; it helps deal with the problem of his presenting his own lab's findings in what he sees as the most effective context without seeming to blow his own horn.

The depersonalization also helps in presenting controversy. The two views around which Darnell organizes his plot are not attributed to anyone, though it would have been possible to find names on which to hang them; each is presented entirely by a depersonalized "view."

> A contrasting view for the origin of divided genes is that . . .
> This view denies an intermediary role in evolution to . . .
> According to this view, the separated DNA segments . . .
> In this scheme, where RNA-RNA splicing is held as basic . . .

It is usual to introduce one's own hypothesis as one of several alternatives, the others of which are then vanquished. But it may be unusually cautious to keep stressing the hypothetical nature of the belief in every sentence. (When Darnell and Doolittle reviewed the work of the field in *PNAS* eight years later, they did give names and citations for the two views. But I would argue that the device of identifying two views worked differently then; in 1978 they were creating two views as a way of putting forward a claim; by 1986 there really were two established views, with extensive literature on each side.)

Although Darnell's presentation is carefully impersonal, in Cricks article, one is immediately aware of a strongly personal, colloquial, spoken voice.

> I have been so rash as to say, more than once, that we might expect between 10 and 100 different enzymes; but that was pure guesswork. The number could be as low as two.

But as one reads on one can hear, not just the voice of Francis Crick, given special privilege as an eminent biologist, but two voices. The flow of the text breaks into a kind of dialogue between the Crick who speculates and the Crick who reports, between the Crick who gives "an overall view of the present position" and the Crick who gives "some general ideas and a few remarks about future work." We can hear two voices in the contrast between the bold statement in

> In a higher organism a gene has, if anything, more nonsense than sense in it.

and the cautious qualification that follows

> These preliminary estimates are necessarily very insecure.

The cautious voice is like a critic restraining the more colorful statements. But it is not just that Crick talks to himself; he talks to others in the field too. There are actual comments from other researchers in the text:

> Gilbert has pointed out to me . . .
> A reasonable guess, as supposed by Tonegawa . . .

Personal communications like these have a special status in the text; they get a name and a comment, while citations of published work just get a footnote number, as in Darnell's article. Another dialogic device is the use of notes that comment on the text, 17 of them, functioning as asides that respond to but are not part of the main thread. So, for instance, this statement is followed by a note qualifying it:

> Thus, one should not invoke some selective advantage occurring only in the future unless this is likely to happen within a time comparable to the time needed to remove the intron (63).
>
> 63. Not all inserts now present need to have a function. For all we know a fair proportion of them may be sitting there, doing nothing, and simply waiting to be excised or deleted.

The sense of back and forth comment is strengthened by the contrast between the style of the text and the even less formal style of the note. Here the anthropomorphism of sitting there is part of a comic style. He replaces the textbook voice and authority of a paper like Darnell's with the sort of license granted to ideal discussion after a conference paper. This is an article with a mug of beer in its hand. Suggestions are thrown out and followed up, half serious guesses are allowed, what is known is ticked off in citations and what isn't known is ticked off in questions. So the result is not as personal to Francis Crick as it seemed from the opening sentence of the article or the more brash speculations; rather, he plays, and allows us to play, the typical molecular geneticist in the audience of these reports, turning over the new findings, looking for something interesting to do.

Crick's article differs from Darnell's more typical review article in several features: Crick's shorter than usual sentences, his less complex syntax, and his preference for cohesion by replacement and substitution rather than by repetition and conjunction. These stylistic differences can be related to differences in rhetorical strategy, if we think of them as indicating the kinds of readers the two articles imply.

Crick's tendency to use short sentences and Darnell's to use long sentences may seem to be matters of personal or editorial taste, but these choices also have rhetorical implications. One can take this sentence from Darnell's article as typical not only of his style but of the style of many reviews.

> For example, although the genetic code is universal (or nearly so) and the machinery for protein synthesis is quite similar in prokaryotes and eukaryotes, the tRNA molecules for specific amino acids—even including initiator tRNA (37, 38), and ribosomal RNAs (rRNA) (35, 36, 39)—bear little resemblance even between

lower eukaryotes and prokaryotes while there is considerable sequence overlap between various eukaryotic tRNA and rRNA molecules.

The reason this sentcnce is so complex is that it both makes a statement and incorporates all the objections that might be raised or attitudes that might be held by readers. The basic statement here is that the tRNA molecules for specific amino acids bear little resemblance between eukaryotes and prokaryotes. This statement is modified by two assertions that broaden the statement:

(1) that even initiator tRNA is included (the readers must expect resemblance here—perhaps because this tRNA codes for a processing function common to all translation), and
(2) that even lower eukaryotes (which the reader might expect to be closer to prokaryotes) are included.

There are also three assertions that put the statement in the context of the current state of disciplinary knowledge, making it more surprising, more worthy of the reader's attention:

(3) the comparison at the end, to the sequence overlap within eukaryotes, of tRNA and rRNA,
(4) the observation at the beginning that the genetic code is universal,
(5) and another observation at the beginning that protein synthesis is similar.

Then in almost every statement there is some qualifying adverb or adjective that makes the statement more cautious:

(or nearly so)
quite similar
little resemblance
considerable sequence overlap

And at the beginning there is a phrase telling us to take this whole thing as just one example of the striking differences between prokaryotes and eukaryotes. Such complex sentences are appropriate to Darnell's task of constructing two clear views of the evidence. In this case, even if one does not know what a particular sentence is saying, one can see the argument has two – and only two – sides. Each piece of evidence is incorporated in a way that acknowledges implicitly the initial skepticism of the implied holders of the other view.

It is not so much the average length of Crick's sentences that makes them seem short compared to Darnell's, as the apparent baldness of some of their assertions and the way they follow each other with only implicit connections.

> 47. Where are split genes to be found? 48. So far, they have only been noticed in eukaryotes. 49. If they were common in prokaryotes (the bacteria and the blue-green algae), they would almost certainly have been discovered earlier. 50. We cannot yet say categorically that they do not occur in prokaryotes but it certainly seems unlikely that they do. 51. They are common in eukaryotic viruses. 52. Indeed, that is where their importance was first realized, but an interesting distinction exists. 53. They have only been found in DNA viruses that occur in the cell nucleus *(2)* or in RNA retroviruses which have a DNA nuclear phase (16). 54. Split genes have not so far been discovered in viruses that exist only in the cytoplasm of a cell. (265; sentence numbers added)

'The short sentences, parallel assertions in similar forms *(they have been noticed, they were common, they are common, they have been found*), seem to cry out for some conjunctions and subordination. But a closer look shows a back-and-forth movement between statements on the distribution of split genes (sentences 48, 51, 53, 54) and statements about the research on them (49, 50, 52). The direction of argument becomes clear only gradually; the point emerges at the end of a series of steps ("All this would suggest that the phenomenon of splicing is correlated with the existence of a nuclear membrane") instead of being given at the beginning, as in the passage from Darnell. The style assumes the readers can put these statements together, can themselves see the relations as they develop.

If we compare the two authors' cohesive devices, we see that Darnell prefers conjunction and repetition, while Crick prefers replacement and substitution (see Halliday and Hasan for terms). The difference in effect is hard to describe, since we have only a hazy understanding of how cohesion works, but it seems that, again, Darnell's prose suggests logical demonstration, while Crick's suggests a more open-ended process. For instance, in the paragraph from which I took the sample sentence about prokaryotes and eukaryotes, Darnell starts most sentences with logical connectors.

> 44. Both prokaryote[s] and eukaryotes existed at least 1.0 to 1.5×10^9 years ago (34, 35), and studies to date provide no evidence of sequential prokaryotic to eukaryotic evolution (35-37). 45. *For example,* although the genetic code is universal (or nearly so) and the machinery for protein synthesis is quite similar in prokaryotes and eukaryotes, the tRNA molecules for specific amino acids—even including initiator tRNA (37, *38),* and

> ribosomal RNA's (rRNA) (35, 36, 39)—bear little resemblance even between lower eukaryotes and prokaryotes while there is considerable sequence overlap between various eukaryotic tRNA and rRNA molecules. 46. *Furthermore*, even in yeasts, which are among the least complex eukaryotic organisms, some tRNAs are formed from a precursor tRNA by the removal of fifteen to twenty nucleotides from the middle of the tRNA sequence (26–28), with subsequent RNA-RNA splicing. 47. *Likewise*, there is little evidence of any overlap between prokaryotes and eukaryotes of primary amino acid sequences even for similar proteins *although*, it must be admitted, there has been very little work done on which such comparisons can be made (37, 40, 41). 48. *Thus*, there is at present no evidence of a "core" or residue of prokaryotic genes that are still present within a now expanded set of eukaryotic genes. (1258; sentence numbers added, connectors in italic)

The whole passage would make some sense even if one didn't know what prokaryotes and eukaryotes were, because every relation is explicitly marked. In sentence 45 the shift to a specific piece of evidence is marked by *for example*. (And we have already seen that this long sentence is full of internal conjunctions.) The next two sentences are marked as continuing the same line of evidence with the conjunctions *furthermore* and *likewise*. The *thus* in the last sentence marks it as ending this line of evidence. (In this case, the argument is complicated by the fact that it is based on negative evidence—demonstrating that certain correlations that should hold if eukaryotes descended from prokaryotes do not in fact hold.)

Besides using conjunctions, Darnell repeats words and phrases, so the reader can identify the same topic in each sentence. The repetition in sentence 44 serves to repeat a topic from the previous paragraph, which then continues through the paragraph:

> 44. sequential prokaryotic to eukaryotic evolution
> 45. little resemblance even between lower eukaryotes and prokaryotes
> 47. little evidence of any overlap between prokaryotes and eukaryotes
> 48. "core" or residue of prokaryotic genes

Such repetition allows the reader to follow the main point through a series of difficult sentences. The constant topic is part of what gives the sense of exhaustive argument to Darnell's style. As we will see, the substitution of a word ("this") for a phrase has a different stylistic effect, and also reflects a different pattern of topics.

Despite this complexity of this paragraph, anyone could sit down and produce an outline of it, or of the whole article. The extremely, even ponderously tight cohesion seems to be designed to lead the reader through the argument, even at the cost of overloading some sentences with embedding and loading each paragraph with conjunctions and repetition. The article could end with Q.E.D.

If close cohesion is characteristic of Darnell's style, Crick's style is characterized by apparent gaps that can only be filled by a reader with some knowledge of the field. I see this loose cohesion as another device for involving the reader in the discussion. The use of lexical replacement and substitution seems to demand more from readers than Darnell's logical conjunctions and repetitions, but it has the effect of including them, implicitly, in the set of intended readers, those familiar with this semantic network. For instance, consider the links between sentences in this passage.

> 120. What is the actual mechanism of splicing? 121. At the moment any ideas must necessarily be largely speculative. 122. One would certainly expect at least one *enzyme* to be involved, if not several. 123. In the case of *tRNA* from *yeast,* an *enzyme* activity has been found by two groups, as was mentioned above, although it has still to be purified (11, 12). 124. It is not completely obvious that such a *mechanism* would require a source of *energy* since two phosphate ester bonds need to be broken whereas one (or possibly two) have to be made. 125. On balance, one would suspect that *energy* might be required if only because the process must be an accurate one. 126. Preliminary evidence indicates that the enzyme appears to need adenosine triphosphate (ATP) (11). (266; repetition and lexical replacement in italic)

One can construct a reader for this article based on the knowledge Crick assumes in connecting one clause to another, one sentence to another. So, for instance, sentences 120 and 122 are held together if one connects mechanism and enzyme Sentences 125 and 126 are linked here only if we connect energy and adenosine triphosphate. Sentences 127 and 128 (not shown here) are linked by a contrast between *tRNA* and *mRNA*. On the basis of repetition we can also provide the connections suggesting a relationship of hypothesis and confirmation between 122 and 123 and an adversative relationship between sentences 124 and 125.

While Darnell maintained a constant topic by repeating key words in each sentence, Crick tends to move from topic to topic through substitution (such as the pronoun "this" substituting for a previous statement).

> 103. Let us now consider in more detail the arrangement of introns and exons. 104. The *first thing* we notice, from the very

> limited experimental data at present available to us, is that a chromosomal gene only produces a single protein (45), whereas a stretch of DNA in a virus may produce more than one protein, depending on which way the primary transcript is spliced. 105. I adopt the attitude that in most case this is because viruses are short of DNA and, by various devices, their limited amount of DNA is made to code for more proteins than would otherwise be possible. 106. We can see this even in proltaryotic viruses . . .

In these sentences, the *first thing* substitutes for the whole statement that is to come, while the *this* in sentences 105 and 106 each refer back to the whole statement of the previous sentence. The effect of substitution is different from that of repealing the whole phrase or a variant on it, as Darnell might do. It is unusual in a scientific article to do this so much, and the use of the device may be part of what makes Crick's style seem so informal. Crick's choices of cohesive devices are consistent with the creation of a sense of discussion, as Darnell's cohesive devices are consistent with his emphasis on the logical resolution of two views.

Melting into the Stream of Knowledge

I have argued that these articles tell stories that try to enlist readers in a particular view of the present and future of the field. Now, almost ten years later (a long time in molecular genetics), we have some idea of how the field did develop. The articles themselves have become parts of other provisional histories of the field. We can see how these and other reviews might have influenced the course of research, both in articles that specifically cite them and more generally in changes in the discourse of the field.

The usual way sociologists gauge the influence of a text is by tracing citations and perhaps examining their contexts (see Swales, "Citation Analysis"; Cozzens). A review is typically cited often, but not for long; in its brief time, it may he read by many more researchers than any experimental report. Darnell's and Crick's articles are both cited a great deal (for instance, in 1980, 39 times for Darnell and 80 for Crick), [9] but random checks of the contexts of these citations suggest they are cited in somewhat different ways. Darnell's article is nearly always cited in relation to one issue, the views of evolution he proposes, while Crick's is cited for a variety of general ideas and specific phrases. I would suggest that these different citation histories result in part from the different stories the two articles present. As we have seen, Darnell organizes a complex series of narratives around one story of how evolution could have occurred. In Crick's article, the underlying story of the shaking up of the research field is not so easily summarized. The styles of the articles may also have something to do with the way they are cited, for Darnell's strategy of presenting two impersonal "views" means he is easily assimilated into one line of the literature, while Crick's personal and informal style means he can make a number of quotable phrases and intriguing remarks.

We can find an example of a typical citation of Darnell in a recent article:

> As has been suggested [cites Darnell, Doolittle, Blake, and Reanney] the split gene organization may have been present in the original ancestral cells, with only the . . . present day prokaryotes and lower eukaryotes having lost their introns under selective pressure. (Michelson et al., 6969)

The issue of whether introns were introduced in eukaryotes or eliminated from prokaryotes has continued to be a focus of research, with new information of various sorts relating to one side or the other of the debate (so, for instance, the Michelson article is interpreting the sequence of the gene they study in terms of this issue). When Darnell is cited for this, he is always cited with one or more of a cluster of other articles. Sometimes, as in his own later articles and Doolittle's (and one they coauthored), his position is presented in contrast to that of the researchers who saw introns as being introduced between prokaryotes and eukaryotes. More often, articles with all sorts of speculations are cited together, just to show that the issue has received considerable attention.

> The problems of the origins of introns, and the possible evolutionary advantage of the split gene organization have been the subject of intense speculation (11, 122–129). Although we have very little chance ever to answer the basic question of whether the split gene organization is the most primitive one and whether present-day bacteria are "streamlined" cells, there are a number of observations that are relevant. . . . (Breathnach and Chambon, 359)

Only rarely is Darnell's article cited for one of its more detailed claims. In another passage from Michelson's article, he is cited for his argument on why split genes are preserved.

> While the original amino acid sequence homology of the common nucleotide binding domain would be eliminated over time as point mutations accumulated, the similarity in the arrangement of exons encoding this domain would be preserved by the relatively

> infrequent occurrence of precise intron loss, at least in the genomes of higher eukaryotes (ref. 30, see below).

But such specific credits are hard to disentangle from the many similar arguments made at the same time. There was a general convergence on evolutionary issues, and in this case the fact of convergence matters more than who said what first. As Crick says about evolution in his review, "I've noticed that this question has an extraordinary fascination for almost everyone concerned with the problem" (269).

Crick's article, too, is most often cited as part of this cluster of discussions of evolutionary issues. Many of the citations of Crick's article in more specialized articles refer to specific claims, but usually these turn out to be the claims of others that he is just evaluating. So, for instance, Wallis, writing in 1980, cites Crick after a reference to "AG, a base sequence found almost universally at the 3' end of an intervening sequence," but this does not imply that the idea was Crick's—it was based on Chambon's work, and Crick just gives a good short account of it. As late as 1984, Crick is cited by Levenson for the statement that the "intervening sequences . . . may be larger than the coding region itself." This is not Crick's own idea, but again he does put this in a striking phrase ("In a higher organism a gene has, if anything, more nonsense than sense in it"). Interestingly, he seems to be frequently cited for the term "exon shuffling" (Darnell is cited for this too). This is Crick's way (I believe he coined the catchy phrase) of describing one of the proposals in an article by Gilbert. Crick is also frequently cited for the proposal that "these genes have evolved from already distinct exons, each coding for a different structural domain." This too is proposed in Gilbert's article. That so many articles cite Crick for this does not necessarily indicate any confusion about the source of the proposal, but just sends the reader to Crick's more evaluative and comparative treatment rather than to Gilbert's rather compressed and sometimes cryptic first statement.

Of course, Crick is also cited for interpretations that, as far as I know, appear for the first time in his review. When Chambon cites his discussion of the mechanism of splicing, it may be because his comment on the ambiguity of the splicing sequence ("an interesting point which is perhaps not immediately obvious") is not made in other reviews. Ahelson takes Crick's guess about the number of enzymes as worth repeating: "Francis Crick estimates (182) that there are more than ten but less than one hundred RNA splicing enzymes, which seems a good guess" (1059). Crick's treatment of possible scenarios for the origin of split genes is also different from that in other reviews of the time.

Several popularizations quote Crick's review, partly because he has a striking way of putting things, and partly because he is Francis Crick.

Marcel Blanc, writing in La Recherche*,* starts by quoting the first two sentences of the article, so that the estimation of "l'un des pionniers de cette discipline moderne" indicates the importance of split genes. Then Blanc presents the area in terms of a controversy between the old guard, like Crick, who try to adapt this discovery within the framework of earlier molecular biology, and somewhat younger researchers like Gilbert who see it as revolutionary.

> In a way that is typical in the history of sciences, the new generation of researchers involved in the discovery are inclined to see it as a revolution, while the attitude of certain pioneers of the discipline, as Francis Crick shows in this *Science* article, is to try to make these "new facts" fit in the old theory, while making some concessions. (My translation)[10]

I do not agree with this way of dividing up the participants, but I can see how Crick's article, by telling a story about the responses of researchers, and by ending, as we have seen, with some cold water, could be useful to a journalist trying to construct his own provisional history of the field.

Reviews may also relate to texts that do not specifically cite them. Taken together, they can make or reflect changes in the discourse of the field. They can do this just be defining problems in general terms, rather than in terms of a specific program of research. For instance, all the research articles are illustrated, usually with electron micrographs, schema, and maps showing the structure of one gene or set of genes. Crick's review article is the first, I believe, to start with a very simplified illustration that shows split genes in general terms, without reference to any particular case. And in the text, he begins by laying out "The Basic Problem," not by relating one particular line of research to others. It is a crucial step in the making of a concept, like Walter Gilbert's naming of the excised portions as "introns" (in "Why Genes in Pieces?") confirming (what had perhaps already been implied) that these sequences were entities to be studied rather than just stuff between the expressed sequences to be studied. Similarly, after the introductory survey, Darnell writes, not about E. coli or adenovirus or rabbit beta-globins or chicken ovalbumin, but about prokaryotes and eukaryotes in general.

Another sign of this process of generalization is Crick's and Darnell's reflection on the basic terms of split genes research. A note in Crick's second sentence points out that, with this discovery, the term *gene* becomes problematic:

> Throughout the article I have deliberately used the word "gene" in a loose sense since at this time any precise definition would be premature.

Then he digresses from his explanation of the problem to discuss Gilbert's new terminology of introns and exons, going on in a note:

> There are two main difficultics. . . . Nevertheless, used judiciously, the two words are undoubtedly useful. I imagine some committee will eventually decide on a wholly logical terminology.

This care with terminology reflects Crick's concern with examining the underpinnings of the theoretical framework. Darnell shows the same awareness by putting every new term or slightly colloquial expression in cautious quotation marks.

One bizarre kind of influence may come simply from giving a name to a notion so that it can be discussed. This explains why Crick's article is sometimes cited for something he explicitly said he wasn't talking about. A year after it, Leslie Orgel and Crick published one articIe, and W. F. Doolittle and Carmine Sapienza another, on the topic of "Selfish DNA." In introducing this topic, Doolittle and Sapienza say that "Dawkins, Crick, and Bodmer have briefly alluded to it" (601). But the mention by Crick in the 1979 article is really very brief. After criticizing the fallacy of "evolutionary foresight," he says, "This problem should *not* be confused with the related phenomenon of a particular stretch of DNA spreading within the genome, the case of 'selfish DNA." But he doesn't talk about what selfish DNA is. Apparently there had been enough informal discussion of this idea in Crick's circle to make it worth addressing in this offhand way, but the idea did not yet have the general currency it has today, which allows the phrase to be used without explanation or citation. This odd use of Crick's article brings us back to the basic difference between his style and Darnell's. It is characteristic of Crick's influence that it should be personal, that a phrase and a suggestive, offhand remark would be cited later; his earlier reviews also contributed a number of striking phrases and broad generalizations to the field. In the same way, it is characteristic of Darnell's work that his influence should be inseparable—and that he should want it to be seen as inseparable—from the influence of other researchers exploring similar ideas at the time. These different kinds of influence reflect two sides of the same strategy of enrollment (Latour 118); the successful researcher must both have allies to carry out the program and have his or her individual contribution to this program recognized. Crick's article illustrates the importance of attribution of credit, while Darnell's illustrates the importance of allies, but both illustrate the same process that is essential to either text having an influence. As I said in introducing Crick and Darnell, the importance of the exception is that it too can be explained in terms of the function of the review article, the use of the past to shape the future.

Recently there have been several retrospective articles that look back at these articles that looked forward; for instance, Gilbert summarizes the research related to his short review in "Genes in Pieces Revisited," and Darnell and Doolittle review research related to their 1978 articles in "Speculations on the Early Course of Evolution." These articles show that much experimental research — sequencing new genes or identifying new biochemical processes or defining a whole new taxonomical kingdom — has pursued just those evolutionary problems laid out in the review articles of 1978 and 1979. But this correlation leaves open the question of whether reviews actually shape future research or whether they just give convenient citations for points of view that are established in the discourse in other, less formal ways. One example will show how the reviews figure as justification for a new line of investigation. P. Senapathy recently published a statistical analysis of available sequences, and he introduces his project with a reference to Darnell and Crick and the various reviews of the late 1970s.

> The origin and function of eukaryotic introns has been an enigma since their discovery and there have been contrasting discussions — whether the introns were introduced when eukaryotic genes evolved from more ancient prokaryotic intron-less genes or whether the primitive single-celled eukaryotes were the most ancient to evolve with introns (5–9). . . . This paper presents a hypothesis for the origin of introns based on statistical analyses of codon distributions in DNA sequences from data banks. . . . (2133)

Now it may be that evolutionary questions are never far from the mind of any biologist. But it may be that a series of reviews by some of the best-known figures in the field made it easier to focus an introduction, like Senapathy's, on these very broad issues, rather than on more narrowly defined issues. There are always plenty of alternative contexts. Whether reviews do shape the field, or merely seem to shape it by reflecting other social interactions, the subsequent development of research along the lines laid out in the reviews shows that they did not just sum up past research; they created a story that continued into the future.

My own aim parallels that of the genre I am describing: I want to enlist other textual analysts in an incomplete project, to get them to do broader and deeper studies of review articles and of what is usually considered "secondary" literature. Traditional sociologists and historians of science have focused on experimental reports as the key scientific texts. This is partly a reflection of the traditional concern with matters of priority and the roles of individuals in discovery — a desire to find the real beginning. But if we are interested in the social construction of science rather than in individual credit, we need to look at some of the other, later textual forms, such as reviews, news articles, textbooks, and popularizations.

Perhaps we look for the original form of each idea because the notion that there is no original form is deeply unsettling. As Max Delbruck remarked in his Nobel Prize address, science involves the dissolution and reconstruction of textual forms.

> While the artist's communication is linked forever with its original form, that of the scientist is modified, amplified, fused with the ideas and results of others, and melts into the stream of knowledge and ideas which forms our culture.

Literary critics have begun to look beyond the "original form" of artistic communications to the "stream of knowledge" of which they are a part. Analysts of scientific texts, too, need to look more closely at the textual forms in which the original communication is modified, amplified, fused, and melted.

NOTES

My thanks to J. E. Darnell, Jr., for his comments on and corrections of an earlier draft, and to Tony Dudley-Evans of the English Language Research Department of the University of Birmingham, and to his postgraduate seminar on genre, for their discussion of an earlier version of this paper.

1. For articles on the review article form, see Eugene Garfield; A. M. Woodward; and J. A. Virgo. For representative handbook and style guide comments, see M. OConnor and F. Peter Woodford; Robert Day; and Janet S. Dodd.

2. This view of reviews is supported by Charles Bazerman's comments (now in *Shaping Written Knowledge)* on how the writing of a review article shaped the thinking of the physicist Arthur Compton.

3. For instance, when I write an article for other researchers within the specialty of social. studies of science (such as "Texts as Knowledge Claims"), I can assume a commitment to relativism, and a methodological interest in showing the variability of accounts. When I review work in the field for readers in other disciplines (as in "Writing Research and the Sociology of Scientific Knowledge"), I have to define relativism and accounts in other terms, and I have to make explicit the assumption that the possibility of variable accounts supports the relativistic interpretation of knowledge, thus raising the question of whether the assumption is correct. And of course I have to show why people who are not immediately involved with this research— such as writing teachers- should invest their time and effort in understanding and applying it.

4. For accounts of the influence of Cricks 1957 review, see Horace Freeland Judson (a huge, readable, *New Yorker*-style account with wonderful interviews), Franklin H. Portugal and Jack S. Cohen (a more academic account), John Gribbin (an interesting example of popularization, but sometimes odd in the slant on details), and, perhaps the best introduction for absolute

beginners, the comic book *DNA for Beginners* (Rosenfeld et al.). Pnina Abir-Am has a detailed critique of the historiography of molecular biology that is relevant to the problem of "influence."

5. This account assumes that the reader knows some textbook facts about genes: that genes are instructions for protein production encoded in the base sequence of DNA molecules; that one strand of the DNA double helix is copied onto complementary single-stranded messenger RNA in a process called transcription; that this mRNA then goes to the ribosomes, where transfer RNA reads each triplet of bases and lines up the appropriate amino acid, forming a protein in a process called translation. These processes were elucidated by research on bacteria (prokaryotes), particularly Escherichia coli, in the 1960s and 1970s. Work on the much more complicated cells with their DNA in nuclei (eukaryotes) was hampered by a number of technical difficulties until new techniques were developed in the mid-1970s.

I have analyzed other texts about split genes in "Making a Discovery," and in an unpublished conference paper comparing two News and Views articles, "Scientific Speculation and Literary Style." Neither of these are historical articles. For popular histories of split genes research, see the article in Scientific American by Pierre Chambon, or the article in La Recherche by Antoine Danchin and Piotr Slonimski. It should be noted, for those who (like me) know science only through the studies of historians and sociologists, that this episode, though dramatic, and though often described as a "revolution," was nothing like the kind of revolution Kuhn describes. RNA splicing could be accounted for within the basic principles of molecular genetics. Nor does it seem to have been a particularly controversial topic; though there were different views about mechanisms, about evolutionary sequence, and other issues, all the researchers seem to have accepted one interpretation of the adenovirus experiments very quickly (what it was like at the symposium I don't know). This is partly because the interpretation put on their results by the CSH and MIT groups enabled other groups, working on hemoglobin, immunoglobin, and ovalbumin, to make sense of results they had been trying to interpret (for a sense of this response, see Chambon, 'The Molecular Biology of the Eukaryotic Genome Is Coming of Age").

6. The plot/story distinction, and its variants, have entered Anglo-American literary criticism through Russian Formalism and the structuralism of Gerard Genette and Roland Barthes. There are many introductions in English, the most useful of which, for me, remains Fredric Jameson's *The Prison-House of Language*. Seymour Chatman's *Story and Discourse* discusses these distinctions in detail, and the first chapter of Peter Brooks' *Reading for the Plot* provides references to more recent studies. Note than E. M. Forster's distinction between plot and story in Aspects of the Novel is different from the distinction I am following.

7. It may well be that Crick's choices of what to ignore, or to restrict to a mention, are significant. When he leaves out most of the work on viruses and immunoglobulin, he presents his omission as a problem of space, but it also means he does not have to give much attention to claims for a "multiple-choice" gene.

8. The one exception to Darnell's impersonality is the mention of W. F. Doolittle's name in the text and a brief discussion of his paper in a note. The note explains why he might be a special case requiring personal recognition:

While this manuscript was being prepared. a note appeared in Nature (London) [272, 581 (1978)] by W F. Doolittle which proposes the same general premise as this article —that splicing and"genes in pieces" is representative of an early phase of cell evolution. Doolittle then goes on to advocate . . .

Darnell gives a name where it is a question of adequately acknowledging the priority of another article. He seems to be bending over backward to give full credit to Doulittle, while also stressing that he himself had the idea independently. (My paper "The Pragmatics of Politeness" comments on such acknowledgements.:)

9. While neither Darnell's article nor Crick's article is among the author's most cited papers, they were both cited a great deal from the time of their publication until 1982, and both have continued to be cited some since then

	Darnell	Crick
1979	16	21
1980	39	80
1981	34	76
1982	29	40
1983	19	24
1984	7	21
1985	13	24
1986	11	12

The first citations take them as the most current review articles, cited with the first reference to splicing, split genes. or introns, so that the reader can catch up ("For review, see . . ."). But they are still cited long after they could have been current, for instance, after Breathnach and Chamhon's long review in 1981, and after masses of new data relevant to the evolution of split genes had been published. This suggests that they both have some appeal that the other, more up-to-date reviews didn't have. It is interesting that most of the earlier citations are in Cell, Nature, PNAS, and Science, the most prestigious places for breaking research in molecular biology. Later citations are more often in review journals and journals of fields outside molecular biology (Crick was even cited in an anthropology review), perhaps suggesting that those at the core of the nucleic acid research soon began citing more recent reviews.

10. "De maniere characteristique dans l'histoire des sciences, la nouvelle generation dechercheurs a lorigine de la decouverte penche pour le bouleversement; landis que l'attitude de certains pionniers de la discipline, comme Francis Crick le montre dans cet article de Science, est d'essayer de faire I-entre les faits nouveaux' dans la theorie ancienne, tout en faisant des concessions" (897).

BIBLIOGRAPHY

Abelson, John. "RNA Processing and the Intervening Sequence Problem." *Annual Review of Biochemistry* 48 (1979): 1035-69.

Abir-Am, Pnina. "Themes, Genres, and Orders of Legitimation in the Consolidation of New Scientific Disciplines: Deconstructing the Historiography of Molecular Biology." *History of Science* 23 (1985): 73-117.

Bazerman, Charles. *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science.* Madison: University of Wisconsin Press, 1988.

Blanc, Marcel. "Une mini-revolution en genetique moleculaire." *La Recherche* 103 (September 1979): 896-98.

Breathnach, Richard, and Pierre Chambon. "Organization and Expression of Eukaryotic Split Genes Coding for Proteins." *Annual Review of Biochemistry* 50 (1981): 349-83.

Brooks, Peter. *Reading for the Plot: Design and Intention in Narrative.* Oxford: Oxford University Press, 1984.

Chambon, Pierre. "The Molecular Biology of the Eukaryotic Genome Is Coming of Age." *Cold Spring Harbor Symposium Proceedings* 42 (1977): 1209-34.

Chambon, Pierre. "Split Genes." *Scientific American* (April 1981): 48–59.

Chatman, Seymour. *Story and Discourse: Narrative Structure in Fiction and Film.* Ithaca: Cornell University Press, 1978.

Cozzens, Susan. "Comparing the Sciences: Citation Context Analysis of Papers from Neuropharmacology and the Sociology of Science." *Social Studies of Science* 15 (1985): 127-53.

Crick, F. H. C. "On Protein Synthesis." *Symposia of the Society for Experimental Biology* 12 (1958): 138-63.

Crick, Francis. "Split Genes and RNA Splicing." *Science* 204 (1979): 264–71.

Crick, F. H. C., Leslie Barnett, S. Brenner, and R. J. Watts-Tobin. "General Nature of the Genetic Code for Proteins." *Nature* 192 (1961): 1227-32. Rpt. in Taylor, ed.

Danchin, Antoine, and Piotr P. Slonimski. "Les genes en morceaux." *La Recherche* 155 (May 1984): 616–26.

Darnell, James E., Jr. "Implications of RNA-RNA Splicing in Evolution of Eukaryotic Cells." *Science* 202 (1978): 1257-60.

Darnell, James E., Jr. "Variety in the Level of Gene Control in Eukaryotic Cells." *Nature* 297 (1982): 365-71.

Darnell, James E., Jr., and W. F. Doolittle. "Speculations on the Early Course of Evolution." *Proceedings of the National Academy of Sciences 83* (1986): 1271-75.

Darnell, James E., Jr., R. Evans, N. Fraser, S. Goldberg, J. Nevins, M. Salditt-Georgieff, H. Schwartz, J. Weber, and E. Ziff. "The Definition of Transcription Units for mRNA." *Cold Spring Harbor Symposium Proceedings* 42 (1977): 515-22.

Day, Robert. *How to Write and Publish a Scientific Paper.* Philadelphia: ISI Press, 1979.

Delbruck, Max. "A Physicist's Renewed Look at Biology: Twenty Years Later (Nobel Lecture)." *Science* 168 (1970): 1312-15.

Dodd, Janet S., ed. *The ACS Stylc Guide: A Manual for Authors and Editors.* Washington: American Chemical Society, 1986.

Doolittle, W. Ford, and Carmen Sapienza. "Selfish Genes, the Phenotype Paradigm and Genome Evolution." *Nature* 284 (1980): 601-3.

Fleck, Ludwik. *The Genesis and Development of a Scientific Fact.* 1935. Trans. Fred Bradley and Thaddeus J. Trenn. Chicago: University of Chicago Press, 1979.

Garfield, Eugene. "Reviewing Review Literature, Part 1. Definitions and Uses of Reviews." *Current Contents* (4 May 1987): 3-8.

Garfield, Eugene. "Reviewing Review Literature, Part 2. The Place of Reviews in the Scientific Literature." *Current Content* (11 May 1987): 3-8.

Gilbert, Walter. "Why Genes in Pieces?" *Nature* 271 (1978): 501.

Gilbert, Walter. "Genes in Pieces Revisited." *Science* 228 (1985): 823-24.

Gribbin, John. *The Search for the Double Helix: Quantum Physics and Life.* London: Corgi, 1985.

Halliday, M. A. K., and Ruqaiya Hasan. *Cohesion in English.* Harlow: Longman, 1976,

Jameson, Fredric. *The Prison-House of Language.* Princeton: Princeton University Press, 1972.

Judson, Horace Freeland. *The Eighth Day of Creation:* The *Makers of the Revolution in Biology.* London: Jonathan Cape, 1979.

Latour, Bruno. *Science in Action: How to Follow Scientists and Engineers Through Society.* Milton Keynes: Open University Press, 1987.

Levenson, Robert, Vincent Racaniello, Lorraine Albritton, and David Housman. "Molecular Cloning of the Mouse Ouabain-Resistance Gene." *Proceedings of the National Academy of Sciences (USA)* 81 (1984): 1489-93.

Marx, Jean L. "Gene Structure: More Surprising Developments." *Science* 199 (1978: 517-18.

Michelson, A. M., C. C. F. Blake, S. T Evans, and S. H. Orkin. "Structure of the Human Phosphoglycerate Kinase Gene and the Intron-Mediated Evolution and Dispersal of the Nucleotide Binding Domain." *Proceedings of the National Academy of Sciences* 82 (1985):6965-69.

Myers, Greg. "Making a Discovery: Narratives of Split Genes." In *Narrative in Culture,* ed. Christopher Nash. London: Routledge and Kegan Paul, 1990.

Myers, Greg. "The Pragmatics of Politeness in Scientific Articles." *Applied Linguistics* 10 (1989): 1-35.

Myers, Greg. "Text as Knowledge Claims: The Social Construction of Two Biologists' Articles." *Social Studies of Science* 15 (1985): 593-630.

Myers, Greg. "Writing Research and the Sociology of Scientific Knowledgc." *College English* 48 (1986):595 610.

O'Connor, M., and F. P. Woodford. *Writing Scientific Papers in English.* Oxford: Elsevier, 1975.

Orgel, L. E.. and F. H. C. Crick. "Selfish DNA: The Ultimate Parasite." *Science* 284 (1980): 604-7.

Portugal, Franklin H., and Jack S. Cohen. A *Century of DNA: A History of the Discovery* of *the Structure and Function of the Genetic Substance.* Cambridge: MIT Press, 1977.

Potter, Jonathan, and Margaret Wetherell. *Discourse and Social Psychology.* Beverly Hills and London: Sage, 1987.

Rogers, John. "Genes in Pieces." *New Scientist* (5 January 1978): 18–20.

Rosenfeld, Israel, Edward Ziff, and Borin Van Loon, *DNA for Beginners.* London: Writers and Readers, 1983.

Sambrook, Joseph. "Adenovirus Amazes at Cold Spring Harbor." *Nature* 268 (1.977):101-4.

Schmeck, Harold M., Jr. " 'Nonsense' in Gene is Prompting New Thoughts On Man's Origin." *The New York Times* (3 November 1981), C1 .

Senapathy, P. "Origin of Eukaryotic Introns: A Hypothesis Based on Codon Distribution Statistics in Genes, and Its Implications." *Proceedings of the National Academy of Sciences* 83 (1986): 2133-37.

Swales, John. *Aspects of Article Introductions.* Birmingham, Eng.: Aston University, 1981.

Swales, John. "Citation Analysis and Discourse Analysis." *Applied Linguistics* 7 (1986): 39–56.

Taylor, J. Herbert, ed. *Selected Papers on Molecular Genetics.* New York and London: Academic Press, 1965.

Virgo, J. A. "The Review Article: Its Characteristics and Problems." *Library Quarterly* 41 (1971): 275.

Wallis, M. "Growth Hormone: Deletions in the Protein and Introns in the Gene." *Science* 284 (1980): 512.

Woodward, A. M. "Review Literature: Characteristics, Sources, and Output in 1972.," *Aslib Proceedings* 26 (1974): 367–76.